# Using deep learning to identify (urban) form and function in satellite imagery - the case of Great Britain

Martin Fleischmann*[1], Daniel Arribas-Bel†[1]

[1]Department of Geography and Planning, University of Liverpool

January 14, 2022

**Summary**

In this paper, we introduce a framework to leverage satellite data through AI to build rich representations of urban form and function. We use a concept of Spatial signatures to characterise predominantly urban environment into data-driven classes based on both form and function composed of a large number of input data sources. Consequently, we explore the ability of Sentinel 2 satellite imagery and state-of-the-art AI models to capture the same classification of space using a single, regularly updated data source, including the conceptual questions of relationship between granular signature geometry and rigid raster grid of satellite data.

**KEYWORDS:** spatial signatures, classification, remote sensing, artificial intelligence

The way we organise cities matters. From walkability (Pafka and Dovey, 2016) to energy efficiency (Ewing and Rong, 2008), the structure of space both reflects and affects the daily life of billions of people. *Spatial signatures* are a concept that encodes the form and function of (urban) environments into a manageable set of distinct, data-driven classes. However, developing signature-based classifications requires a large amount of diverse data updated slowly over time. The question then stands: How to roll the classification back and forward, acknowledging varied update cycles of underlying data? The resource that could help and is continuously up to date is satellite imagery. However, all the information of relevance contained in an image is in unstructured form. To provide structure and decode such information, we train a neural network to capture spatial signatures from Sentinel 2 satellite imagery. In effect, this potentially allows turning the static classification into a temporal one. In this paper, we discuss the potential and challenges of such an approach, what can be captured (and what cannot), and where is the current limit of satellite data when it comes to an understanding of the form and function of urban environments as well as the relationship between the signature geometry and raster data being conceptualised as a grid of a certain pixel size.

The building blocks that make up cities, the activities and agents that inhabit them, and the structure that supports them can be spatially arranged in many ways. Activities and agents can be conceptualised as urban *function*, while the structure as urban *form*; two aspects of an environment worth quantifying and understanding in their own right but, more importantly, in tandem. However, the study of form and function is fragmented, scattered across many disciplines and levels of the policy hierarchy. That means that while individual aspects of understanding are well covered - economic forces shaping cities by economists (Ahlfeldt and Pietrostefani, 2019), transportation networks by planners (Gil, 2014), or public space by urban designers and architects (Khirfan, 2011) - the connection is not always present. There is thus a clear need for detailed, consistent and scalable evidence on urban form and function.

This paper relies on the concept of “spatial signatures”, a characterisation of space based on form and function designed to understand urban environments (Arribas-Bel and Fleischmann, 2022). Spatial signatures exhaustively divide geographical space into distinct classes based on its appearance (form) and how it is used (function). This division is based on two underlying concepts – enclosed tessellation

* m.fleischmann@liverpool.ac.uk
† d.arribas-bel@liverpool.ac.uk

as a basic spatial unit and quantitative characterisation as a way of capturing form and function. The enclosed tessellation is an indivisible (any subdivision would result in a unit that is incapable of capturing the nature of urban form and function), internally consistent (reflects only a single signature type), and exhaustive (covering the entirety of the study area) spatial unit derived from a set of natural or built barriers or delimiters (as rivers, railways, or roads) and anchors such as building footprints. Each of the resulting enclosed tessellation cells is then characterised using urban morphometrics (to capture form) and by linking additional functional data to this new geometry, where the latter includes information on population, points of interest, land use, land cover and other characteristics.

Due to their dependency on a multitude of data sources that are being updated at a variable and often slow rate (e.g., ten years for the case of census-based data) signatures cannot be easily updated with frequency. The challenge that stands ahead is thus enabling spatial signatures as a temporal classification. That means being able to use available data to derive signatures for points in a time predating the recent data and, importantly, being able to capture their state going into the future periodically. With predominantly vector-based data linked to various statistical and administrative boundaries (e.g., census output areas or uniform grids), depending on the occasional collection and release of data, being a significant component of the function part of signatures, the method needs to decouple itself from these data sources. At the same time, it needs to preserve its open nature, i.e., being based on transparent procedures as well as openly available data.

One possible pathway towards this goal comes from remote sensing and satellite imagery. While staying in the realm of open data, two substantial resources are available - the Sentinel mission of the European Space Agency and the Landsat mission of NASA/USGS. Of these, two stand out: the Sentinel 2 satellite mission offers visible and near-infrared bands at 10 meters per pixel resolution, while Landsat 8 provides similar bands (among others) at 30 meters resolution. Both sources are of specific interest as they able to capture urban structure at a similar scale that that which spatial signatures focus on (e.g., spatial pattern composed of a contiguous area of multiple, often thousands, of enclosed tessellation cells). Moreover, there is suggestive evidence these resources are able to capture urban patterns of an analogous scale. Examples include the detection of different Local Climate Zones (Taubenböck et al., 2020), land use and land cover patterns (Georganos et al., 2018), or urban structural types (Huck et al., 2011).

We explore this pathway using the Sentinel 2 imagery within a deep convolutional neural network (CNN) trained to predict each pixel's spatial signature type across Great Britain. This exploration aims not only to develop the optimal model with the highest accuracy but also to understand the relationship between granular geometry of signatures and conceptualisation of space into a grid and rectangular input chips (a small piece of the grid) of raster data. With Sentinel 2 being relatively coarse in terms of the spatial resolution compared to some commercial products that can offer a resolution of up to 30 cm per pixel (Maxar, 2022), there are not only technical questions of the CNN architecture, but also geographical ones related to a Modifiable Areal Unit Problem (Openshaw, 1983) and the ability of chips of a certain size to capture the nature of each signature type. While larger chips are naturally better at capturing more information, they often cover multiple signature types. On the other hand, smaller chips that may fit better within individual signature types may not carry enough information to result in a reliable prediction.

This paper discusses this exploratory work and outlines our empirical experiments. We talk through the options we face in the selection of a neural network, its training, and relevant parameters. We identify the crucial variables and investigate their effects on the performance of the model both in terms of accuracy and complexity, leading to a longer training time. Since spatial signatures are composed not only of urban form that can be seen by a naked eye but also of urban function that is harder to "see", the paper tries to find the limit of satellite imagery and understand what can be captured from space using primarily bands falling into the visible spectrum and what cannot.

While traditional data sources such as census, population estimates, or crowdsourced points of interest (like those available via the OpenStreetMap project) are extremely valuable in understanding urban environments, and hence deriving spatial signatures, they are limited in their ability to provide temporal granularity. On the other hand, satellite imagery comes with frequent and consistent updates but does not by itself provide enough information to build spatial signatures from scratch. However, developing a CNN that uses spatial signatures based on traditional data and learns to predict them from satellite imagery is possible and opens a way towards developing up-to-date temporal classifications of form and function in urban environments. While such a pursuit is not without both technical and conceptual challenges, it has the potential to significantly expand the applicability of spatial signatures as a method of understanding cities and minimise their current data dependency.

**Acknowledgements**

The authors kindly acknowledge funding by the UK's Economic and Social Research Council through the project "Learning an urban grammar from satellite data through AI", project reference ES/T005238/1.

**References**

Ahlfeldt, G. M. and Pietrostefani, E. (2019). The economic effects of density: A synthesis. *Journal of Urban Economics*, 111:93–107.

Arribas-Bel, D. and Fleischmann, M. Spatial Signatures - Understanding (urban) spaces through form and function (2022). Mimeo.

Ewing, R., Rong, F., 2008. The impact of urban form on US residential energy use. *Housing policy debate* 19, 1–30. https://doi.org/10/fgf7q8

Georganos, S., Grippa, T., Vanhuysse, S., Lennert, M., Shimoni, M., and Wolff, E. (2018). Very high resolution object-based land use–land cover urban classification using extreme gradient boosting. *IEEE geoscience and remote sensing letters*, 15(4):607–611.

Gil, J., 2014. Analyzing the configuration of multimodal urban networks. *Geographical Analysis* 46, 368–391. https://doi.org/10.1111/gean.12062

Huck, A.; Hese, S.; Banzhaf, E. Delineating parameters for object-based urban structure mapping in Santiago de Chile using QuickBird data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci*. 2011, 38, 6.

Khirfan, L., 2011. Understanding the links between inherited built forms and urban design: Athens and Alexandria as case studies. *Urban Morphology* 15, 39–53.

Maxar, 2022. Optical Imagery. URL https://www.maxar.com/products/optical-imagery (accessed 11.1.22).

Openshaw, S., 1983. The modifiable areal unit problem, *Concepts and techniques in modern geography*. Geo Books.

Pafka, E., Dovey, K., 2016. Permeability and interface catchment: measuring and mapping walkable access. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability* 10, 150–162. https://doi.org/10.1080/17549175.2016.1220413

Taubenböck, H., Debray, H., Qiu, C., Schmitt, M., Wang, Y., Zhu, X.X., 2020. Seven city types representing morphologic configurations of cities across the globe. *Cities* 105, 102814. https://doi.org/10/gg2jv4

## Biographies

Martin Fleischmann is a Research Associate in the Geographic Data Science Lab at the University of Liverpool. His research interest lies in urban morphology and geographic data science focusing on quantitative analysis of urban form, remote sensing and AI.
Dani Arribas-Bel is interested in computers, cities, and data. He is a senior lecturer in Geographic Data Science at the the University of Liverpool, and Deputy Programme Director for Urban Analytics and ESRC fellow at the Alan Turing Institute.